國立政治大學 應用數學系

碩士 學位論文

霍奇排名之理論分析

Theoretic Aspect of HodgeRank

碩士班學生：陳名秀　撰
指導教授：蔡炎龍　博士

中華民國 105 年 6 月 30 日

國 立 政 治 大 學 應 用 數 學 系

陳名秀君所撰之碩士學位論文

霍奇排名之理論分析
Theoretic Aspect of HodgeRank

業經本委員會審議通過
論文考試委員會委員：

指導教授：

系主任：

中華民國 105 年 6 月 30 日

致謝

經過兩年的努力，終於將自己的論文完成並通過口試。非常感謝蔡炎龍老師指導我時所付出的時間與心力。在這段日子裡，非常幸運地可以遇到蔡炎龍老師讓我對於論文有正確的方向及嚴謹審慎的架構與邏輯。若是沒有老師的耐心的指導，我的論文是不會這麼順利完成。

對於口試委員也非常的感激，能夠給予我許多寶貴的意見，以及對於論文的建議和指點出我論文的問題。

最後也謝謝同儕們與朋友們和家人們給予的鼓勵、支持和幫助。

陳名秀　謹誌于

國立政治大學應用數學系碩士班

中華民國 105 年 6 月

中文摘要

霍奇排名是在近幾年才運用在排名的一種方法。在大多數現在的資料庫中，資料庫很龐大，有些甚至會需要網路連結，而且很多會有資料不完整或是資料不平衡的狀況。我們選擇用霍奇排名這種排名方法來處理可能會遇到的這些困擾。

這篇論文主要目的是想用運用基本的線性代數來研究霍奇排名和推導組合霍奇理論。

關鍵字：霍奇理論, 霍奇排名, 組合霍奇理論

Abstract

HodgeRank is a method of ranking that is new in recent years. In most of modern datasets, the amount of data is very large, some also need the internet connection, and plenty of them have the feature that incomplete or imbalanced. We use the method of HodgeRank to deal with these difficulties.

This thesis is primary using elementary linear algebra to survey HodgeRank and deduce the combinatorial Hodge Theorem.

Keywords: HodgeRank, combinatorial Hodge theorem, combinatorial Hodge decomposition,

Contents

List of Figures

Chapter 1

Introduction

The technique of solving ranking problems has many applications. For instance, recommendation systems of restaurants, movies, books, and so on. It is difficult to apply traditional ranking methods to many modern datasets, since these datasets usually are incomplete and imbalanced.

In [6], Jiang, Lim, Yao, and Ye developed a method called HodgeRank to deal with these difficulties.

We use a simple example to elaborate the idea of HodgeRank. Suppose that we have four candidates A, B, C, D to rank, and we have the following pairwise comparisons:

- A is better than C by 1 point,
- B is better than C by 1 point,
- C is better than D by 1 point, and
- D is better than B by 2 points.

We can represent the relations by a graph, as shown in the Figure 1.1.

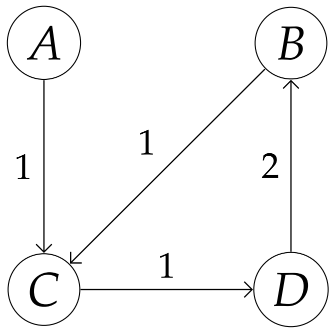

Figure 1.1: Draw the result on the graph.

The same relations can also be presented by a matrix

$$Y = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 2 \\ 1 & 1 & 0 & -1 \\ 0 & -2 & 1 & 0 \end{pmatrix} \end{array}.$$

Note that Y is a skew-symmetric matrix. The goal is to find a score function

$$s \colon \{A, B, C, D\} \to \mathbb{R}.$$

If this can be done, then we get a global ranking of A, B, C, D.

Usually the task is not trivial, for we might have some kind of inconsistency in our datasets. In our example, we can see

$$B > C > D > B$$

which is a contradiction.

The HodgeRank approach is to apply so called *combinatorial Hodge decomposition* to the matrix Y

$$Y = X + X_H + X_T,$$

and the matrix X will give us the score function in desire.

In this thesis, we survey the HodgeRank and use elementary linear algebra to deduce the combinatorial Hodge theorem. The structure of this thesis is organized as follows. In chapter 2, we explain the main application we have in mind, which is the pairwise ranking problems; In chapter 3, we prove all important theorems we need from elementary linear algebra; In chapter 4 we introduce the combinatorial Hodge theory; In chapter 5, we apply the combinatorial Hodge theory to graphs; and finally we give the conclusion in the final Chapter.

Chapter 2

Pairwise Ranking

2.1 Pairwise Ranking Problems

Let $V = \{v_1, v_2, \ldots, v_n\}$ be a collection of *alternatives* (or *canditates*). We would like to give the set of alternatives a global ranking. For instance, V can be a collection of sport teams, movies, books, or students from a class, and so on. Suppose that we can do some pairwise ranking. For example, team A is better than team B by 5 points; book C is 2.3 points better than D. However, these datasets usually unbalanced, incomplete, even inconsistence. In this chapter, we will describe the type of ranking problems we consider and discuss the difficulties we will encounter. For more details we refer to [2, 3, 9–14].

2.2 Introduction to HodgeRank

In this thesis, we use the method HodgeRank [6] to deal with pairwise ranking. To elaborate the idea of HodgeRank, we give an example. Support we have four sport teams $V = \{A, B, C, D\}$. Four games was held and the results are as follows

- A beats C for 1 point,
- B beats C by 1 point,
- C beats D by 1 point, and

- D beats B by 2 points.

Then we can present the results on a graph, as figure 2.1 shows.

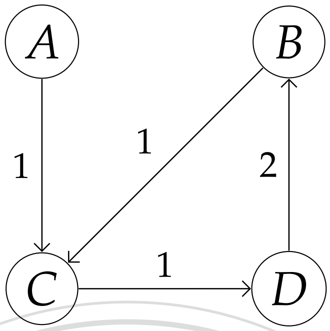

Figure 2.1: The graph of pairwise ranking.

One problem is that we can't intuitively determine which team is the best for the results are *inconsistent*. For example, C beats D and D beats B. However, B beats C by 1 points.

We can also present the results by the corresponding skew-symmetric matrix,

$$Y = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 2 \\ 1 & 1 & 0 & -1 \\ 0 & -2 & 1 & 0 \end{pmatrix} \end{array}.$$

Thus a pairwise ranking can be represented by a skew-symmetric matrix. The observation gives us the following definition.

Definition 2.2.1 (Pairwise Ranking)**.** A *pairwise ranking* of n alternatives is an $n \times n$ skew-symmetric matrix.

Many different methods of dealing pairwise ranking have been used in many fields, like psychology, management science, social choice theory, and statistics [1, 4, 5, 7, 14]. Our goal is to get a global ranking from a pairwise ranking. For our example, that means that we want to find a function

$$s\colon \{A, B, C, D\} \to \mathbb{R}.$$

We call the function s a *score function*.

Applying so called the combinatorial Hodge decomposition to Y, we get

$$Y = X + X_H + X_T,$$

where X is exactly what we want:

$$X = \begin{pmatrix} 0 & -1.33 & -1 & -0.67 \\ 1.33 & 0 & 0.33 & 0.66 \\ 1 & -0.33 & 0 & 0.33 \\ 0.67 & -0.66 & -0.33 & 0 \end{pmatrix}.$$

Remember that we have an inconsistent pairwise ranking, but now X is consistent. Set a score for A then we can get the function s in desire. For example, let $s(A) = 0.75$. From the matrix X, we get the score function

$$s(B) = -0.58,$$
$$s(C) = -0.25, \text{ and}$$
$$s(D) = 0.08.$$

2.3 Possible Applications of Pairwise Rankings

Suppose we want to rank n items $v_1, v_2, \ldots, v_n$, the key of our approach is to model our problem into a pairwise ranking. We give some possible examples as follows.

2.3.1 Ranking Students in a Class

Many pairwise rankings raised from grading or voting from some voters. We denote all voters by $\Lambda = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$. For example, the peer assessment in MOOCs, students will give other's homework scores online with the criterion. For most people, rating 100 alternatives at a time is harder than rating only 2 alternatives at a time. In fact, it has been observed that most people can rank only between 5 and 9 alternatives at a time. [8] And pairwise ranking has fewer values missing. Take the Netflix problem in the [6] for an example, there is almost 99% of its values missing but only 0.22% of the pairwise comparison values are missing. Moreover, in certain things, such as a badminton tournament, only pairwise comparison is possible.

2.3.2 Movie Recommendation Systems

Suppose the movie recommendation system wants to rank the movies in 2016, candidates are all movies that be released in the movie theaters in 2016 and voters are all viewers that had ever seen some of these movies. The dataset is all the movies, and then the movie recommendation system wants to globally rank all the movies. Now for simplicity, suppose we only want to rank these three movies, "Me Before You," "Independence Day: Resurgence," and "The Conjuring 2," which we denote by M_1, M_2 and M_3. Each time the voters answer that "preferring M_1 or M_2?", "preferring M_2 or M_3?", or "preferring M_1 or M_3"?" But there may be a case: if voter A thinks M_1 is better than M_2, voter B prefers M_2 to M_3 and voter C votes M_3 is better than M_1. It results an inconsistent cyclic preference relation

$$M_1 > M_2 > M_3 > M_1. \tag{2.3.1}$$

2.3.3 Ranking Sports Team

We can use a graph to represent the dataset: the vertices are all candidates we have to rank and the edges means that the end points of each edge have been compared. In other words, if we want to rank football teams, the vertex set is all the football teams and if two teams have a match to each other then there is an edge between these two vertices. In pairwise comparison, we can assign the edge a direction and weight, as edge flows, on a diagraph. There are many different ways to define the weight, for instance, we can let the weight to represent the scores difference, the score of winning candidate subtracts the score of the losing candidate, and use an arrow to represent the direction from the winning candidate to the losing one.

2.4 Pairwise Ranking from Voting

Let $V = \{v_1, v_2, \ldots, v_n\}$ be the set of candidates to be ranked and $\Lambda = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be the set of voters. For each $\alpha \in \Lambda$, the pairwise comparisons of voter α is a skew-symmetric matrix $X^{\alpha} \in \mathbb{R}^{n \times n}$. For each ordered paired $(i, j) \in V \times V$, we have $X^{\alpha}_{ij} = -X^{\alpha}_{ji}$. In matrix X^{α}, X^{α}_{ij} means that how much the degree of αth voter preferring jth candidate to i-th candidate. If αth voter does not compare ith and jth candidate, X^{α}_{ij} is a missing value and we let X^{α}_{ij} be zero. The diagonal entries on the corresponding matrix X are all zero since when we compare someone and itself, there is no one of them is better than the other. There are several ways to define the degree of preference, like the score difference. For a game, if one wins by a majority of 1, it is clearly that there is another one losses for 1 point. So we can construct a corresponding skew-symmetric matrix and we can apply the combinatorial Hodge Theory to obtain a global ranking.

2.5 Problems of Pairwise Ranking

As we have seen, many problems can be transferred into a pairwise ranking. However, there are three major problems in raw pairwise ranking datasets. The pairwise rankings raised from real-world problems usually are

inconsistent: As we have seen in 2.3.1.

incomplete: In most cases, not all voters will vote (grade) all candidates.

imbalanced: As in the movie recommendation system, one movie might be scored by many voter, but another might be scored by just one or two persons.

Use the method of HodgeRank, we can solve all these problems.

Chapter 3

Background

3.1 Some Simple Facts in Linear Algebra

In this chapter, the definition of inner product space is an inner product space over the filed $\mathbb{R}$ or $\mathbb{C}$ (Hermitian product space).

Lemma 3.1.1. Let V be a finite dimensional inner product space over $\mathbb{R}$, and let $g\colon V \to \mathbb{R}$ be a linear transformation.Then there exists a unique vector $v' \in V$ such that $g(v) = \langle v, v' \rangle$ for all $v \in V$.

Theorem 3.1.2. Let V, W be two finite dimensional inner product space. Let $T\colon V \to W$ be a linear transformation. Then there is unique linear transformation $T^*\colon W \to V$ such that

$$\langle T(v), w \rangle_W = \langle v, T^*(w) \rangle_V.$$

We call T^* the *adjoint* linear transformation of T.

Proof. Let $w \in W$, and define $g\colon W \to \mathbb{R}$ by $\langle T(v), w \rangle_W$ for all $T(v) \in W$.

Let $T(v_1), T(v_2) \in W$ and $c \in \mathbb{R}$.

$$\begin{aligned} g(cv_1 + v_2) &= \langle T(cv_1 + v_2), w\rangle_W \\ &= \langle cT(v_1) + T(v_2), w\rangle_W \\ &= c\langle T(v_1), w\rangle_W + \langle T(v_2), w\rangle_W \\ &= cg(v_1) + g(v_2). \end{aligned}$$

Hence, g is linear.

By Lemma 3.1.1 , there exists $v' \in V$ such that $g(v) = \langle v, v'\rangle_V$, i.e $\langle T(v), w\rangle_W = \langle v, v'\rangle_V$.

Define $T^*\colon W \to V$ by $T^*(w) = v'$, we have $\langle T(v), w\rangle_W = \langle v, T^*(w)\rangle_V$.

To show T^* is linear. Let $w_1, w_2 \in W$ and $c \in \mathbb{R}$.For all $v \in V$,

$$\begin{aligned} \langle v, T^*(cw_1 + w_2)\rangle_V &= \langle T(v), cw_1 + w_2\rangle_W \\ &= c\langle T(v), w_1\rangle_W + \langle T(v), w_2\rangle_W \\ &= c\langle v, T^*(w_1)\rangle_V + \langle v, T^*(w_2)\rangle_V \\ &= \langle v, cT^*(w_1) + T^*(w_2)\rangle_V. \end{aligned}$$

Since v is arbitrary , $T^*(cw_1 + w_2) = cT^*(w_1) + T^*(w_2)$.

Finally, we show T^* is unique.

Given $U\colon W \to V$ is linear and $\langle T(v), w\rangle_W = \langle v, U(w)\rangle_V$, for all $v \in V$ and $w \in W$.

Then $\langle v, T^*(w)\rangle_V = \langle T(v), w\rangle_W = \langle v, U(w)\rangle_V$, for all $v \in V$ and $w \in W$.

So we get $T^* = U$. □

Theorem 3.1.3. Let V, W be two finite dimensional inner product space. Let $T\colon V \to W$ be a linear transformation. If β, γ be arbitrary orthonormal bases for V, W, respectively, and $A = [T]_\beta^\gamma$, then

$$[T^*]_\gamma^\beta = A^*.$$

Proof. Let $\beta = \{v_1, v_2, ..., v_n\}$ and $\gamma = \{w_1, w_2, ..., w_m\}$. Then

$$A = \begin{pmatrix} \langle T(v_1), w_1\rangle_W & \langle T(v_2), w_1\rangle_W & \cdots & \langle T(v_n), w_1\rangle_W \\ \langle T(v_1), w_2\rangle_W & \langle T(v_2), w_2\rangle_W & \cdots & \langle T(v_n), w_2\rangle_W \\ \vdots & \vdots & \ddots & \vdots \\ \langle T(v_1), w_m\rangle_W & \langle T(v_2), w_m\rangle_W & \cdots & \langle T(v_n), w_m\rangle_W \end{pmatrix},$$

and we get

$$A^* = \begin{pmatrix} \langle T(v_1), w_1\rangle_W & \langle T(v_1), w_2\rangle_W & \cdots & \langle T(v_1), w_m\rangle_W \\ \langle T(v_2), w_1\rangle_W & \langle T(v_2), w_2\rangle_W & \cdots & \langle T(v_2), w_m\rangle_W \\ \vdots & \vdots & \ddots & \vdots \\ \langle T(v_n), w_1\rangle_W & \langle T(v_n), w_2\rangle_W & \cdots & \langle T(v_n), w_m\rangle_W \end{pmatrix}.$$

Let $B = [T^*]_\gamma^\beta$, then

$$\begin{aligned} B &= \begin{pmatrix} \langle T^*(w_1), v_1\rangle_V & \langle T^*(w_2), v_1\rangle_V & \cdots & \langle T^*(w_m), v_1\rangle_V \\ \langle T^*(w_1), v_2\rangle_V & \langle T^*(w_2), v_2\rangle_V & \cdots & \langle T^*(w_m), v_2\rangle_V \\ \vdots & \vdots & \ddots & \vdots \\ \langle T^*(w_1), v_n\rangle_V & \langle T^*(w_2), v_n\rangle_V & \cdots & \langle T^*(w_m), v_n\rangle_V \end{pmatrix} \\ &= \begin{pmatrix} \langle v_1, T^*(w_1)\rangle_V & \langle v_1, T^*(w_2)\rangle_V & \cdots & \langle v_1, T^*(w_m)\rangle_V \\ \langle v_2, T^*(w_1)\rangle_V & \langle v_2, T^*(w_2)\rangle_V & \cdots & \langle v_2, T^*(w_m)\rangle_V \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, T^*(w_1)\rangle_V & \langle v_n, T^*(w_2)\rangle_V & \cdots & \langle v_n, T^*(w_m)\rangle_V \end{pmatrix} \\ &= \begin{pmatrix} \langle T(v_1), w_1\rangle_W & \langle T(v_1), w_2\rangle_W & \cdots & \langle T(v_1), w_m\rangle_W \\ \langle T(v_2), w_1\rangle_W & \langle T(v_2), w_2\rangle_W & \cdots & \langle T(v_2), w_m\rangle_W \\ \vdots & \vdots & \ddots & \vdots \\ \langle T(v_n), w_1\rangle_W & \langle T(v_n), w_2\rangle_W & \cdots & \langle T(v_n), w_m\rangle_W \end{pmatrix} \\ &= A^*. \end{aligned}$$

$\square$

Theorem 3.1.4. Let V, W be two finite dimensional inner product space. Let $T\colon V \to W$ be a linear transformation. Then

$$V = \ker(T) \oplus \operatorname{im}(T^*),$$

and

$$W = \operatorname{im}(T) \oplus \ker(T^*).$$

Proof. We just prove $V = \ker(T) \oplus \operatorname{im}(T^*)$, for the arguments for $W = \operatorname{im}(T) \oplus \ker(T^*)$ are exactly the same.

Since we have $V = \operatorname{im}(T^*) \oplus \operatorname{im}(T^*)^\perp$. We only need to show that $\ker(T) = \operatorname{im}(T^*)^\perp$.

Let $x \in \ker(T)$. We have $T(x) = 0$. Then

$$\langle T(x), w \rangle_W = 0, \text{ for all } w \in W.$$

The definition of T^* gives us

$$\langle x, T^*(w) \rangle_V = 0, \text{ for all } w \in W.$$

We conclude that $x \in \operatorname{im}(T^*)^\perp$. The proof in the reverse direction is similar. $\square$

Lemma 3.1.5. Let V, W be two finite dimensional inner product space. Let $T\colon V \to W$ be a linear transformation. Then

1. $\ker(T^*) = \operatorname{im}(T)^\perp$,and
2. $\ker(T^*T) = \ker(T)$.

Proof. 1. Let $x \in \ker(T^*)$. We have $T^*(x) = 0$. Then

$$\langle v, T^*(x) \rangle_V = 0, \text{ for all } v \in V.$$

The definition of T^* gives us

$$\langle T(v), x\rangle_W = 0, \text{ for all } v \in V.$$

We conclude that $x \in \mathrm{im}(T)^\perp$.

So we get $ker(T^*) \subseteq \mathrm{im}(T)^\perp$.

And it is similar to proof $\mathrm{im}(T)^\perp \subseteq ker(T^*)$.

Beside, we can prove it using the proof in Theorem3.1.4 by $(T^*)^* = T$.

2. Clearly, $\ker(T) \subseteq \ker(T^*T)$.

   Given $x \in \ker(T^*T)$, $T^*T(x) = 0$.

   It impies

   $$T(x) \in \ker(T^*) = \mathrm{im}(T)^\perp.$$

   But we have

   $$T(x) \in \mathrm{im}(T).$$

   So $T(x) = 0$. Then $x \in \ker(T)$.

   Therefore, $\ker(T^*T) = \ker(T)$.

□

3.2 The First Isomorphism Theorem

Theorem 3.2.1. (The First Isomorphism Theorem). Let $T\colon V \to W$ be a linear transformation between two vector spaces over field $\mathbb{R}$. Then

$$V/\ker(T) = \mathrm{im}(T).$$

Proof. Define $\phi\colon V/\ker(T) \to \mathrm{im}(T)$; that is, $\phi(v + \ker(T)) = T(v)$ for all $v \in V$.

For all $v \in V$, there exists

$$\phi(v + \ker(T)) = T(v) \in \operatorname{im}(T).$$

And for all $v_1 + \ker(T), v_2 + \ker(T) \in V + \ker(T)$ with $v_1 + \ker(T) = v_2 + \ker(T)$, it implies

$$v_1 - v_2 \in \ker(T).$$

Then

$$T(v_1 - v_2) = 0.$$

So we have

$$T(v_1) = T(v_2).$$

Therefore, ϕ is well-defined.

Second, we show that ϕ is linear. For all $\phi(v_1 + \ker(T)), \phi(v_2 + \ker(T)) \in \operatorname{im}(T)$ and $c \in \mathbb{R}$,

$$\begin{aligned}\phi(v_1 + \ker(T) + v_2 + \ker(T)) &= \phi((v_1 + v_2) + \ker(T)) \\ &= T(v_1 + v_2) = T(v_1) + T(v_2).\end{aligned}$$

And

$$\begin{aligned}\phi(c(v + \ker(T))) &= \phi(cv + ker(T)) \\ &= T(cv) = cT(v) \\ &= c\phi(v + ker(T)).\end{aligned}$$

So ϕ is linear.

Then, for all $\phi(v_1 + \ker(T)), \phi(v_2 + \ker(T)) \in \operatorname{im}(T)$ with $\phi(v_1 + \ker(T)) = \phi(v_2 + \ker(T))$, we have $T(v_1) = T(v_2)$. So

$$T(v_1) - T(v_2) = 0,$$

and we get

$$T(v_1 - v_2) = 0,$$

which implies

$$v_1 - v_2 \in \ker(T).$$

Then we obtain

$$v_1 + \ker(T) = v_2 + \ker(T).$$

Hence, ϕ is one-to-one.

Last, for all $w \in T(v)$, there exists $v \in V$ such that $T(v) = w$. We get

$$\phi(v + \ker(T)) = T(v) = w.$$

Therefore, ϕ is onto. □

Chapter 4

Combinatorial Hodge Theory

In this thesis, we only focus on combinatorial Hodge theory. More information of classical Hodge theory we refer to [15, 16].

4.1 Cochain Complex

Let $(C_\bullet, \partial_\bullet)$ be a cochain complex of inner product spaces:

$$\cdots \xrightarrow{\partial_{-1}} C_0 \xrightarrow{\partial_0} C_1 \xrightarrow{\partial_1} C_2 \xrightarrow{\partial_2} \cdots$$

such that ∂_k's are linear transformations and $\partial_{k+1}\partial_k = 0$.

Let ∂_k^* be the adjoint of the linear transformation ∂_k. That is, ∂_k^* is the unique linear transformation from C_{k+1} to C_k such that

$$\langle \partial_k(x), y \rangle_{k+1} = \langle x, \partial_k^*(y) \rangle_k,$$

for all $x \in C_k$ and $y \in C_{k+1}$. We now have the following maps:

$$\cdots \longrightarrow C_{k-1} \underset{\partial_{k-1}^*}{\overset{\partial_{k-1}}{\rightleftarrows}} C_k \underset{\partial_k^*}{\overset{\partial_k}{\rightleftarrows}} C_{k+1} \longrightarrow \cdots.$$

Definition 4.1.1. Let $(C_\bullet, \partial_\bullet)$ be a cochain complex of inner product spaces. Define the *k-th combinatorial Laplacian* $\Delta_k \colon C_k \to C_k$ by

$$\Delta_k = \partial_{k-1}\partial_{k-1}^* + \partial_k^*\partial_k.$$

An element u in C_k is *harmonic* if $\Delta_k(u) = 0$.

4.2 Combinatorial Hodge Theory

Lemma 4.2.1. Let $(C_\bullet, \partial_\bullet)$ be a cochain complex of inner product spaces.

1. $C_k = \operatorname{im}(\partial_{k-1}) \oplus \ker(\partial_{k-1}^*) = \operatorname{im}(\partial_k^*) \oplus \ker(\partial_k)$

It is not hard to prove the following theorem.

Theorem 4.2.2 (Combinatorial Hodge Theorem)**.** Let $(C_\bullet, \partial_\bullet)$ be a cochain complex of inner product spaces. Then we have

1. $H^k(C) = \ker(\partial_k)/\operatorname{im}(\partial_{k-1}) \simeq \ker(\Delta_k)$,
2. $C_k = \operatorname{im}(\partial_{k-1}) \oplus \ker(\Delta_k) \oplus \operatorname{im}(\partial_k^*)$, and
3. $\ker(\Delta_k) = \ker(\partial_k) \cap \ker(\partial_{k-1}^*)$.

We call (2) the *combinatorial Hodge decomposition* on C_k.

Proof. We first prove (3).
Clearly, $\ker(\partial_k) \cap \ker(\partial_{k-1}^*) \subseteq \ker(\Delta_k)$.
Let $x \in \ker(\Delta_k) = \ker(\partial_{k-1}\partial_{k-1}^* + \partial_k^*\partial_k)$, which means $\partial_{k-1}\partial_{k-1}^*(x) + \partial_k^*\partial_k(x) = 0$.
Then

$$\partial_{k-1}\partial_{k-1}^*(x) = -\partial_k^*\partial_k(x).$$

Multiplying ∂_k ,

$$\partial_k\partial_{k-1}\partial_{k-1}^*(x) = -\partial_k\partial_k^*\partial_k(x) = 0.$$

We get $\partial_k^*\partial_k(x) \in \ker(\partial_k)$, but we have $\partial_k^*\partial_k(x) \in im(\partial_k^*) = \ker(\partial_k)^\perp$.

So $\partial_k^*\partial_k(x) = 0$.

We have $x \in \ker(\partial_k^*\partial_k) = \ker(\partial_k)$.

Multiplying ∂_{k-1} ,

$$\partial_{k-1}^*\partial_{k-1}\partial_{k-1}^*(x) = -\partial_{k-1}^*\partial_k^*\partial_k(x) = 0.$$

We get $\partial_{k-1}\partial_{k-1}^*(x) \in \ker(\partial_{k-1}^*)$, but we have $\partial_{k-1}\partial_{k-1}^*(x) \in \mathrm{im}(\partial_{k-1}) = \ker(\partial_{k-1}^*)^\perp$.

So $\partial_{k-1}\partial_{k-1}^*(x) = 0$.

We have $x \in \ker(\partial_{k-1}\partial_{k-1}^*) = \ker(\partial_{k-1}^*)$.

Then $x \in \ker(\partial_k) \cap \ker(\partial_{k-1}^*)$.

We obtaine $\ker(\partial_k) \cap \ker(\partial_{k-1}^*) = \ker(\Delta_k)$.

Second,we prove (2).

$\partial_k \cdot \partial_{k-1} = 0$, then $(\partial_k \cdot \partial_{k-1})^* = \partial_{k-1}^* \cdot \partial_k^* = 0$.

We obtain $\mathrm{im}(\partial_k^*) \subseteq \ker(\partial_{k-1}^*)$.

$$\begin{aligned}\ker(\partial_{k-1}^*) &= C_k \cap \ker(\partial_{k-1}^*)\\ &= [\mathrm{im}(\partial_k^*) \oplus \ker(\partial_k)] \cap \ker(\partial_{k-1}^*)\\ &= [\mathrm{im}(\partial_k) \cap \ker(\partial_{k-1}^*)] \oplus [\ker(\partial_k) \cap \ker(\partial_{k-1}^*)]\\ &= \mathrm{im}(\partial_k^*) \oplus [\ker(\partial_k) \cap \ker(\partial_{k-1}^*)]\end{aligned}$$

Now,

$$\begin{aligned}C_k &= \mathrm{im}(\partial_{k-1}) \oplus \ker(\partial_{k-1}^*)\\ &= \mathrm{im}(\partial_{k-1}) \oplus \mathrm{im}(\partial_k^*) \oplus [\ker(\partial_k) \cap \ker(\partial_{k-1}^*)].\end{aligned}$$

And by (3), $\ker(\partial_k) \cap \ker(\partial_{k-1}^*) = \ker(\Delta_k)$.

Therefore,

$$\begin{aligned}
C_k &= \mathrm{im}(\partial_{k-1}) \oplus \ker(\partial_{k-1}^*) \\
&= \mathrm{im}(\partial_{k-1}) \oplus \mathrm{im}(\partial_k^*) \oplus [\ker(\partial_k) \cap \ker(\partial_{k-1}^*)] \\
&= \mathrm{im}(\partial_{k-1}) \oplus \mathrm{im}(\partial_k^*) \oplus \ker(\Delta_k).
\end{aligned}$$

Last, we prove (1) .

Since $\partial_k \cdot \partial_{k-1} = 0$, we get $\mathrm{im}(\partial_{k-1}) \subseteq \ker(\partial_k)$.

Let $\phi\colon C \to \mathrm{im}(\partial_{k-1})^{\perp}$ be the projection onto the subspace $\mathrm{im}(\partial_{k-1})^{\perp}$.

$C = \mathrm{im}(\partial_{k-1})^{\perp} \oplus \mathrm{im}(\partial_{k-1})$.

So $x = \phi(x) + (1-\phi)(x)$, for all $x \in C$.

And

$$\begin{aligned}
\ker(\partial_k) &= C \cap \ker(\partial_k) \\
&= [\mathrm{im}(\partial_{k-1})^{\perp} \oplus \mathrm{im}(\partial_{k-1})] \cap \ker(\partial_k) \\
&= [\mathrm{im}(\partial_{k-1})^{\perp} \cap \ker(\partial_k)] \oplus [\mathrm{im}(\partial_{k-1}) \cap \ker(\partial_k)] \\
&= [\mathrm{im}(\partial_{k-1})^{\perp} \cap \ker(\partial_k)] \oplus \mathrm{im}(\partial_{k-1}).
\end{aligned}$$

We can restrict ϕ to the subspace $\ker(\partial_k)$, denoted ϕ_k .

Then

$$x = \phi_k(x) + (1-\phi_k)(x), \text{ for all } x \in \ker(\partial_k) .$$

ϕ is surjective , so is ϕ_k.

Then $\mathrm{im}(\phi_k) = \ker(\partial_k) \cap \mathrm{im}(\partial_{k-1})^{\perp}$.

By the first isomorphism theorem,

$$\ker(\partial_k) / \ker(\phi_k) \simeq \mathrm{im}(\phi_k) = \ker(\partial_k) \cap \mathrm{im}(\partial_{k-1})^{\perp}.$$

And

$$\forall x \in \ker(\phi_k), \phi_k(x) = 0. \Rightarrow x \in \operatorname{im}(\partial_{k-1}),$$

also same in the reverse. So we get $\ker(\phi_k) = \operatorname{im}(\partial_{k-1})$.

Then

$$\ker(\partial_k) / \operatorname{im}(\partial_{k-1}) \simeq \ker(\partial_k) \cap \operatorname{im}(\partial_{k-1})^{\perp}.$$

By (3), $\ker(\Delta_k) = \ker(\partial_k) \cap \ker(\partial_{k-1}^*)$, and (1) in Lemma 3.1.5 ,$\operatorname{im}(\partial_{k-1})^{\perp} = \ker(\partial_{k-1}^*)$, we finally get

$$\ker(\partial_k) / \operatorname{im}(\partial_{k-1}) \simeq \ker(\partial_k) \cap \ker(\partial_{k-1}^*) = \ker(\Delta_k).$$

□

4.3 Finding the Best Solution

Let $(C_\bullet, \partial_\bullet)$ be a cochain complex of inner product spaces:

$$C_0 \underset{\partial_0^*}{\overset{\partial_0}{\rightleftarrows}} C_1 \underset{\partial_1^*}{\overset{\partial_1}{\rightleftarrows}} C_2.$$

such that for $k = 0, 1$, ∂_k are linear transformations and ∂_k^* be the adjoint of the linear transformation ∂_k. That is, ∂_0^* is the unique linear transformation from C_1 to C_0 such that

$$\langle \partial_0(x), y \rangle_1 = \langle x, \partial_0^*(y) \rangle_0, \text{ for all } x \in C_0 \text{ and } y \in C_1.$$

And ∂_1^* is the unique linear transformation from C_2 to C_1 such that

$$\langle \partial_1(x), y \rangle_2 = \langle x, \partial_1^*(y) \rangle_1, \text{ for all } x \in C_1 \text{ and } y \in C_2.$$

Fix y in C_1, we want to find

$$\min_{x \in C_0} \| \partial_0 x - y \| .$$

That is, let $V = \{\partial_0 x \mid x \in C_0\}$ and given $y \in C_1$, we want to find the unique $\partial_0 \overline{x} \in V$, where $\overline{x}$ is in C_0, which is closest to y.

We know that $\partial_0 \overline{x} - y \in V^{\perp}$. So for all $x \in C_0$,

$$\langle \partial_0 \overline{x} - y, \partial_0 x \rangle_1 = 0.$$

We get

$$\langle \partial_0^* (\partial_0 \overline{x} - y), x \rangle_0 = 0, \text{ for all } x \in C_0.$$

It implies

$$\partial_0^* (\partial_0 \overline{x} - y) = 0.$$

Then

$$\partial_0^* \partial_0 \overline{x} - \partial_0^* y = 0.$$

Therefore,

$$\overline{x} = (\partial_0^* \partial_0)^{-1} \partial_0 y.$$

Chapter 5

Combinatorial Hodge Theory on Graphs

5.1 Graphs and Functions on Graphs

In this section, we introduce some terminologies in graph theory. The basic definitions are referred to [17]. For representing our data, we introduce special functions called flows.

Definition 5.1.1. A *graph* G is a triple consisting of a vertex set V, an edge set E, and a relation that associates with each edge two vertices (not necessarily distinct) called its endpoints. We write $G = (V, E)$ to present a graph G with vertex set V and edge set E.
Note that the vertex set and the edge set are finite in general. The vertex set V having n elements means that G have n vertices, denoted by $|G| = n$; that is $V = \{v_1, v_2, \ldots, v_n\}$ and writing $v_i v_j$ for an edge with its endpoints $v_i, v_j \in V$.

Definition 5.1.2. A *loop* is an edge whose endpoints are equal and *multiple edges* are edges having the same pair of endpoints.

A *simple* graph is a graph having no loops or multiple edges.

A graph $\overline{G} = (\overline{V}, \overline{E})$ is called a *subgraph* of $G = (V, E)$ if $\overline{V} \subseteq V$ and $\overline{E} \subseteq E$.

Moreover, if the vertices of a simple graph are pairwise adjacent, it is a *complete* graph. And a *clique* is a set of pairwise adjacent vertices in a graph. That is, a complete graph is each distinct pair of vertices (v_i, v_j) in V, the edge $v_i v_j$ is in E and a clique of G is a nonempty complete subgraph of G. A k-clique of G is a clique of G that containing k vertices.

For an example, figure 5.1 is a simple graph because it does not have loop and multiple edges. The vertex set is $\{A, B, C, D\}$ and the edge set is $\{AD, BC, CD, BD\}$. There is no edge between vertex A and vertex B, so it isn't a complete graph.

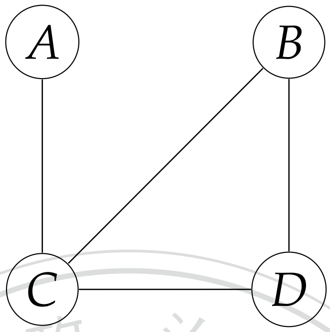

Figure 5.1: A simple graph.

Definition 5.1.3. A *path* is a simple graph whose vertices can be ordered so that two vertices are adjacent if they are consecutive in the list. A v_i, v_j-path, $v_i, v_j \in V$, means that we can travel from v_i to v_j and the vertices in the path are not repeated.

A *cycle* is a graph with an equal number of vertices and edges whose vertices can be placed around a circle so that two vertices are adjacent if and only if they appear consecutively along the circle. We can think that a cycle is a path that the start vertex and the end vertex are the same.

On figure 5.1, there is a cycle: $B - C - D - B$; there is an A, B-path: $A - D - C - B$. There is another A, B-path, $A - D - B$.

Definition 5.1.4. A *directed graph* G (or *diagraph*) is a triple consisting of a vertex set V, an edge set E, and a function assigning each edge an ordered pair of vertices. In the ordered pair (v_i, v_j) in V, the first vertex is called the tail of the edge and the second is the head. And when drawing a diagraph, we draw an arrow from its tail to its head. The vertex set is written by $V = v_1, v_2, \ldots, v_n$ and an edge in E is written as $\overrightarrow{v_i v_j}$ by its tail v_i and head v_j. But in our pairwise ranking problems, all is diagraph. So we only simply write an edge as $v_i v_j$ by its tail v_i and head v_j.

Definition 5.1.5. An *edge flow* X on a simple graph $G = (V, E)$ is a real valued function on the edge set E and $X \colon V \times V \to \mathbb{R}$ satisfies

$$\begin{cases} X(v_i v_j) = -X(v_j v_i), \text{for all } v_i v_j \in E \\ X(v_i v_j) = 0, \text{for all } v_i v_j \notin E. \end{cases}$$

Let $C_1 = \{$ all edge flows $\}$.

We define X is zero for all pairs of vertices that are not adjacent, and $X(v_i v_i) = 0$ since we have defined edges does not self-adjacent; if there had self-adjacent edges, there are multiple edges and G is not a simple graph.

The edge flow function can be represented by a skew-symmetric matrix $[X_{ij}]$ by $X_{ij} = X(v_i, v_j)$, and we still define $X_{ij} = 0$ if $v_i, v_j \notin E$. If there is a $n \times n$ skew-symmetric matrix and a graph G with $|G| = n$, we can define an edge flow of G by letting $X(v_i v_j) = X_{ij}$. So the set of edge flows on G is ono-to-one corresponding to the set of $n \times n$ skew-symmetric matrices, that is satisfying

$$\{X \in M_{n \times n}(\mathbb{R}) | X^T = -X \text{ and } X_{ij} = 0 \text{ if } v_i v_j \notin E\}.$$

From now on, it is not important to distinguish edges flows and skew-symmetric matrices. The reason is we can induce one of them if we have another. As figure 5.2, there is a graph with 4 vertices whose edge flows can be corresponding to the following 4×4 skew-symmetric matrix, and vice versa.

$$Y = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 2 \\ 1 & 1 & 0 & -1 \\ 0 & -2 & 1 & 0 \end{pmatrix} \end{array}.$$

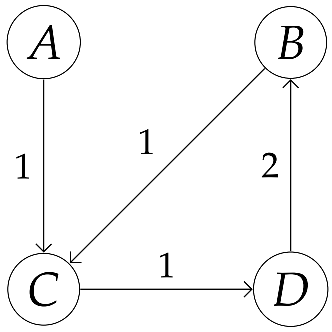

Figure 5.2: A graph with 4 vertices.

Define

$$\binom{V}{k}$$

be the set of k-elements subset of V.

It is obvious that the edge set

$$E \subseteq \binom{V}{2}.$$

Definition 5.1.6. Let $V = \{v_1, v_2, \ldots, v_n\}$ be the set of candidates to be ranked and $\Lambda = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be the set of voters. Define the *weight* function $\omega\colon \Lambda \times V \times V$ by

$$\omega_{ij}^{\alpha} = \omega(\alpha, i, j) = \begin{cases} 1 & \text{, if voter } \alpha \text{ made a comparison for } v_i, v_j \in V \\ 0 & \text{, if voter } \alpha \text{ didn't make a comparison for } v_i, v_j \in V. \end{cases}$$

The weight matrix W is defined by

$$w_{ij} = \sum_{\alpha \in \Lambda} \omega_{ij}^{\alpha}.$$

It is clear that W is a symmetric matrix whose entries are nonnegative values.

In a simple graph $G = (V, E)$, we can find the sets of the collections of 1-clique, 2-cliques and 3-cliques respectively. Obviously, the collection of 1-clique is just the vertex set V. The collection of 2-cliques is the edge set E:

$$E = \left\{ v_i v_j \middle| \{v_i, v_j\} \in \binom{V}{2}, \omega_{ij} > 0 \right\},$$

where $\omega_{ij} = \sum_\alpha \omega_{ij}^\alpha$. The collection of 3-cliques is the set T:

$$T = \left\{ v_i v_j v_k \middle| \{v_i, v_j, v_k\} \in \binom{V}{3}, \text{ and } v_i v_j, v_j v_k, v_k v_i \in E \right\}. \tag{5.1.1}$$

The 3-cliques on the graph is triangles, so we denote T. The set of k-cliques, denoted $K_k(G)$, is defined by

$$\{v_{i_1}, \ldots, v_{i_k}\} \in K_k(G) \Leftrightarrow v_{i_p} v_{i_q} \in E, \forall 1 \leq p < q \leq k.$$

We get

$$K_1(G) = V, K_2(G) = E \text{ and } K_3(G) = T.$$

On page 23 figure 5.1, the triple (B,C,D) is a 3-cliques and the only one.

Definition 5.1.7. A *score function* $s \in \mathbb{R}^n$ is a real-valued function on the set of all candidates; that is,

$$s \colon V \to \mathbb{R}.$$

Let $C_0 = \{\text{all score functions } s\}$.

A score function assign every candidate a score : $s(v_i) = s_i$, for all $v_i \in V$. Every score function induces a global ranking.

Definition 5.1.8. The *combinatorial gradient* or simply *gradient*, denoted grad, is a mapping on the vertices to an edge flow. Define grad: $C_0 \to C_1$ by

$$\text{For all } s \in C_0,\ (\operatorname{grad} s)(v_i v_j) = s_j - s_i, \text{ for all } v_i, v_j \in V.$$

It is obvious that gradient is an edge flow.

Theorem 5.1.9. Gradient flow is independent with path. In other words, if X is an gradient flow, then $X_{ij} = s_j - s_i$ for $v_i, v_j \in V$ for some s is only depending on v_i, v_j no matter what v_i, v_j-path is. In other words, gradient flow is independent with path.

Proof. Let $v_1 - v_2 - \cdots - v_k$ be a path on a graph G with $|G| = n$, and X be a gradient flow. Then the flow of X along the path is

$$X(v_1v_2) + X(v_2v_3) + \cdots + X(v_{k-2}v_{k-1}) + X(v_{k-1}v_k)$$

Since a gradient flow is $X(v_iv_j) = s_j - s_i$ for some s, we get

$$\begin{aligned}&X(v_1v_2) + X(v_2v_3) + \cdots + X(v_{k-2}v_{k-1}) + X(v_{k-1}v_k)\\ =&(s_2 - s_1) + (s_3 - s_2) + \cdots + (s_{k-1} - s_{k-2}) + (s_k - s_{k-1})\\ =&s_k - s_1\end{aligned}$$

□

If we get a globally consistent(see Definition 5.1.14) pairwise ranking X, it is easy to determine a score function s (up to adding a scalar) by solving $\operatorname{grad}(s) = X$. And we can obtain a global ranking of candidates. For example,

Theorem 5.1.10. A gradient is a linear transformation.

Proof. Let $\operatorname{grad}\colon C_0 \to C_1$ be a gradient flow by

$$\text{For all } s \in C_0,\ (\operatorname{grad} s)(v_iv_j) = s_j - s_i, \text{ for all } v_i, v_j \in V.$$

Given $s, t \in C_0$, $(\operatorname{grad} s)(v_iv_j) = s_j - s_i$ and $(\operatorname{grad} t)(v_iv_j) = t_j - t_i$.

For all $v_i, v_j \in V$, we have

$$
\begin{aligned}
\operatorname{grad}(s+t)(v_i v_j) &= (s+t)_j - (s+t)_i \\
&= (s_j + t_j) - (s_i + t_i) \\
&= s_j - s_i + t_j - t_i \\
&= (\operatorname{grad} s)(v_i v_j) + (\operatorname{grad} t)(v_i v_j).
\end{aligned}
$$

And for all $c \in \mathbb{R}$,

$$
\begin{aligned}
\operatorname{grad}(cs)(v_i v_j) &= (cs)_j - (cs)_i \\
&= cs_j - cs_i \\
&= c(s_j - s_i) \\
&= c(\operatorname{grad} s)(v_i v_j).
\end{aligned}
$$

□

Definition 5.1.11. The *triangular flow* on $G = (V, E)$ is the function $\Phi\colon V \times V \times V \to \mathbb{R}$ that satisfies

$$
\begin{aligned}
\Phi_{ijk} &= \Phi(v_i v_j v_k) = \Phi(v_j v_k v_i) = \Phi(v_k v_i v_j) \\
&= -\Phi(v_j v_i v_k) = -\Phi(v_i v_k v_j) = -\Phi(v_k v_j v_i),
\end{aligned}
$$

We define $C_2 = \{\text{all triangle flows}\}$.

Definition 5.1.12. Let X be an edge flow and T in page 26(5.1.1). A *curl* is a map from edge flows to triangle flows. Define curl: $C_1 \to C_2$ by

$$
\text{For all } X \in C_1,\ (\operatorname{curl} X)(v_i v_j v_k) = \begin{cases} X_{ij} + X_{jk} + X_{ki}, \text{if } v_i v_j v_k \in T \\ 0, \text{if } v_i v_j v_k \notin T. \end{cases}
$$

Theorem 5.1.13. A curl is a linear transformation.

Proof. Let $\Phi\colon C_1 \to C_2$ be a curl by

$$\text{For all } X \in C_1,\ (\text{curl } X)(v_iv_jv_k) = \begin{cases} X_{ij} + X_{jk} + X_{ki}, \text{if } v_iv_jv_k \in T \\ 0, \text{if } v_iv_jv_k \notin T. \end{cases}$$

Given $A, B \in C_1$ and for all $v_i, v_j, v_k \in \begin{pmatrix} V \\ 3 \end{pmatrix}$,

If $v_iv_jv_k \in T$,

$$\begin{aligned} \text{curl}(A+B)(v_iv_jv_k) &= (A+B)_{ij} + (A+B)_{jk} + (A+B)_{ki} \\ &= A_{ij} + B_{ij} + A_{jk} + B_{jk} + A_{ki} + B_{ki} \\ &= (A_{ij} + A_{jk} + A_{ki}) + (B_{ij} + B_{jk} + B_{ki}) \\ &= \text{curl}(A)(v_iv_jv_k) + \text{curl}(B)(v_iv_jv_k). \end{aligned}$$

If $v_iv_jv_k \notin T$,

$$\begin{aligned} &\text{curl}(A+B)(v_iv_jv_k) = 0, \\ &\text{curl}(A)(v_iv_jv_k) = 0, \text{ and} \\ &\text{curl}(B)(v_iv_jv_k) = 0. \end{aligned}$$

So, we get

$$\text{curl}(A+B)(v_iv_jv_k) = \text{curl}(A)(v_iv_jv_k) + \text{curl}(B)(v_iv_jv_k).$$

And for all $c \in \mathbb{R}$,

$$\begin{aligned} \text{curl}(cA)(v_iv_jv_k) &= (cA)_{ij} + (cA)_{jk} + (cA)_{ki} \\ &= cA_{ij} + cA_{jk} + cA_{ki} \\ &= c(A_{ij} + A_{jk} + A_{ki}) \\ &= c\,\text{curl}(A)(v_iv_jv_k). \end{aligned}$$

$\square$

Definition 5.1.14. Let $X\colon V \times V \to \mathbb{R}$ be a edge flow on a pairwise comparison graph $G = (V, E)$.

1. X is called consistent on $v_i v_j v_k \in T$ if it is curl-free on $v_i v_j v_k$. That is,

$$(\operatorname{curl} X)(v_i v_j v_k) = X_{ij} + X_{jk} + X_{ki} = 0.$$

2. X is called globally consistent if it is a gradient of a score function. That is,

$$X = \operatorname{grad} s \text{ for some } s\colon V \to \mathbb{R}.$$

3. X is called locally consistent or triangular consistent if it is curl-free on every triangle in T.

4. X is called a cyclic ranking if it contains any inconsistencies. That is,

$$\text{there exists } v_i, v_j, \ldots, v_p, v_q \in V \text{ such that } X_{ij} + X_{jk} + \cdots + X_{pq} + X_{qi} \neq 0$$

Global consistency implies local consistence but the inverse may not be true. There is may be a case that is globally consistent but not locally consistent, as figure 5.3 shown. There is a closed path, $A \to B \to C \to D \to A$, adding all edge flows along the path is $X_{AB} + X_{BC} + X_{CD} + X_{DE} + X_{EA} = 1 + 1 + 1 + 1 + 1 = 5 \neq 0$. We get a nonzero net weight along the path, so figure 5.3 is globally inconsistent. On the other hand, we look at the only triangle in the graph, $A \to B \to C \to A$. Since $\operatorname{curl}(A, B, C) = X_{AB} + X_{BC} + X_{CA} = 1 + 1 - 2 = 0$, figure 5.3 is locally consistent.

Lemma 5.1.15.

$$\operatorname{curl} \circ \operatorname{grad} = 0$$

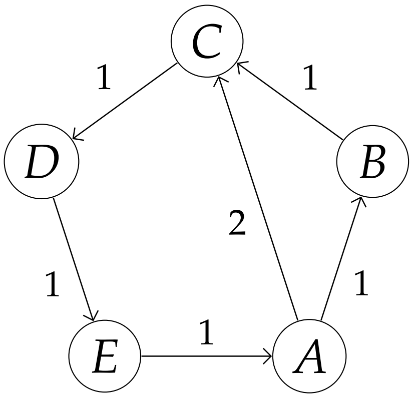

Figure 5.3: An example of globally inconsistent but locally consisitent.

Proof. Let $X = \operatorname{grad}(s)$ for some $s \in \mathbb{R}^n$, and $v_i v_j v_k \in T$. Then

$$\begin{aligned} \operatorname{curl}(X)_{ijk} &= X_{ij} + X_{jk} + X_{ki} \\ &= (s_j - s_i) + (s_k - s_j) + (s_i - s_k) \\ &= 0 \end{aligned}$$

□

From the lemma, we know that global consistency implies local consistency. Then a curl-free on a complete graph must be a gradient.

Definition 5.1.16. A *divergence* $\operatorname{div} \in \mathbb{R}^n$ is a real-valued function on the set of all candidates; that is, $\operatorname{div}: V \to \mathbb{R}$. Define $\operatorname{div}: C_1 \to \mathbb{R}^n$ by

$$\text{for all } X \in C_1,\ \operatorname{div} X(v_i) = \sum_{v_j, (v_i v_j \in E)} X_{ij}.$$

And if $\operatorname{div} X(v_i) = 0$ on each vertex $v_i \in V$, the graph G is called divergence-free.

We also take page 25 figure 5.2 for an example:

$$\operatorname{div}(A) = -1, \operatorname{div}(B) = 1, \operatorname{div}(C) = 1, \operatorname{div}(D) = -1,$$

and figure 5.2 is not divergence-free.

Definition 5.1.17. Let X be an edge flow on the graph G. The flow X is a *harmonic* flow if it both curl-free and divergence-free.

5.2 Matrices

We can use skew-symmetric matrices to represent edge flows, gradient flows, harmonic flows, and curl flows.

Let A be the set of all skew-symmetric matrices:

$$A = \{X \in M_{n\times n}(\mathbb{R}) \mid X^T = -X\}$$

Define the set of gradient matrices:

$$M_G = \{X \in A \mid X_{ij} = s_j - s_i \text{ for some } s \in \mathbb{R}^n\}, \tag{5.2.1}$$

and the set of triangular consistent matrices :

$$M_T = \{X \in A \mid X_{ij} + X_{jk} + X_{ki} = 0 \text{ for all } v_i v_j v_k \in T\}.$$

By lemma 5.1.15, every $X \in M_G$ is in M_T; that is, $M_G \subseteq M_T$. With the usual inner product $\langle \cdot, \cdot \rangle$ in vector space $M_{n\times n}(\mathbb{R})$:

$$\langle A, B \rangle = \operatorname{tr}(A^* B) = \sum_{i,j} a_{ij} b_{ij} \text{ for all } A, B \in M_{n\times n}(\mathbb{R}),$$

we let M_H the orthogonal complement of M_G in M_T; that is,

$$M_H = M_T \cap M_G^{\perp}.$$

It is equivalent to

$$M_T = M_G \oplus M_H. \tag{5.2.2}$$

The elements in M_H is harmonic, which are curl-free and divergence-free. Since M_G and M_T are subspaces of A, we can get

$$
\begin{aligned}
A &= M_G \oplus M_G^{\perp} \\
A &= M_T \oplus M_T^{\perp}
\end{aligned}
\tag{5.2.3}
$$

So we can obtain

$$
\begin{aligned}
A &= M_T \oplus M_T^{\perp} \text{ by equation}(5.2.3) \\
&= M_G \oplus M_H \oplus M_T^{\perp} \text{ by equation}(5.2.2)
\end{aligned}
\tag{5.2.4}
$$

5.3 Inner Product

We define the weighted inner product in C_1 by

$$
\langle X, Y \rangle_W = \sum_{i,j} \omega_{ij} x_{ij} y_{ij}, \tag{5.3.1}
$$

for all edge flows X, Y in C_1. And equation(5.2.4) is still hold.

C_1 is all the set of all $n \times n$ skew-symmetric matrices, the inner product in C_1 defined in (5.3.1) is referred to [6]. But there is a little problem, we need to be careful of it.

If we have a pairwise ranking problem, we obtained the weighted matrix and define the inner product in C_1 in the function (5.3.1). If we give the weight matrix W and $A \in C_1$ as following:

$$
W = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & -1 & -1 & 0 \end{pmatrix}. \tag{5.3.2}
$$

Then

$$\langle A, A \rangle_W = \sum_{i,j} \omega_{ij} a_{ij}^2 = 0$$

but A is not $\boldsymbol{O}$. It is not a inner product.

We try to give some constraints to let it be a inner product. First, all candidates must be voted at least once. So all our graphs are connected. But it is clear that the graph related the matrix $X = \boldsymbol{O}$ is not a connected graph. So the set of all weighted adjacency matrices induced by connected graphs is not a vector space. So we consider the set

$$S = \text{span}\{X \in M_{n\times n}(\mathbb{R}) \mid X \text{ is a weighted adjacency matrix induced by a connected graph}\}.$$

We want (5.3.1) defined in S to be a inner product. But if we take

$$B = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix}.$$

Clearly, B and C are skew-symmetric matrices.

$$B + C = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & -1 & -1 & 0 \end{pmatrix} = A$$

A is in S. And it is the same as the example in (5.3.2). We still check that it is not a inner product in S. We failed to let the definition in (5.3.1) be a inner product in the subspace of the set of all skew-symmetric matrices.

Therefore, we need to give more constraints to get the subspace of C_1 and let our definition in (5.3.1) be a inner product .

5.4 Find the Global Ranking

Our approach to ranking is to minimize a weighted sum of pairwise loss of a global ranking. Our problem is ordinary least squares:

$$\min_{X \in M_G} \sum_{\alpha,i,j} \omega_{ij}^{\alpha}(x_{ij} - y_{ij}^{\alpha})^2,$$

,where M_G is the set that is defined in (5.2.1) and ω_{ij}^{α} is defined in definition 5.1.6.

The optimization problem can be written as

$$\min_{X \in M_G} \| X - Y \|_{2,w}^2 = \min_{X \in M_G} \sum_{i,j} \omega_{ij}(x_{ij} - y_{ij})^2, \tag{5.4.1}$$

where $\omega_{ij} = \sum_{\alpha} \omega_{ij}^{\alpha}$ and y_{ij} is the average scores; that is, $y_{ij} = \sum_{\alpha} \omega_{ij}^{\alpha} y_{ij}^{\alpha} / \sum_{\alpha} \omega_{ij}^{\alpha}$.

5.5 Applying Combinatorial Hodge Theory

If we represent functions on vertices by n vectors, edge flows by $n \times n$ skew-symmetric matrices, and triangular flows by $n \times n \times n$ skew-symmetric matrices. That is,

$C_0 = \{s \in \mathbb{R}^n\}$,

$C_1 = \{X \in M_{n\times n}(\mathbb{R}) \mid X_{ij} = -X_{ij}\}$, and

$C_2 = \{\Phi \in M_{n\times n\times n}(\mathbb{R}) \mid \Phi_{ijk} = \Phi_{jki} = \Phi_{kij} = -\Phi_{jik} = -\Phi_{ikj} = -\Phi_{kji}\}$.

Then, by the theorem 4.2.2, we have

$$\mathrm{im}(\partial_0) = \mathrm{im}(\mathrm{grad}) = M_G,$$

$$\ker(\partial_1) = \ker(\mathrm{curl}) = M_T,$$

$$\ker(\partial_0^*) = \ker(\mathrm{div}) = M_G^{\perp}, \text{ and}$$

$$\mathrm{im}(\partial_1^*) = \mathrm{im}(\mathrm{curl}^*) = M_T^{\perp}.$$

By definition 4.1.1: $\Delta_k = \partial_{k-1}\partial_{k-1}^* + \partial_k^*\partial_k$, and in particular

$$\Delta_0 = \partial_0^*\partial_0 = -\operatorname{div}\circ\operatorname{grad},$$
$$\Delta_1 = \partial_0\partial_0^* + \partial_1^*\partial_1 = \operatorname{curl}^*\circ\operatorname{curl} - \operatorname{grad}\circ\operatorname{div}.$$

Then the optimal solution can be written to

$$\min_{X\in M_G} \parallel X - Y \parallel_{2,w}^2 = \min_{s\in C_0} \parallel \partial_0 s - Y \parallel_{2,w}^2 = \min_{s\in C_0} \parallel \operatorname{grad} s - Y \parallel_{2,w}^2 \tag{5.5.1}$$

Theorem 5.5.1. The solution of 5.5.1 satisfy

$$\Delta_0 s = -\operatorname{div} Y,$$

and the minimum solution is

$$s* = -\Delta_0^\dagger \operatorname{div} Y,$$

where the $\Delta_0^\dagger$ is the Moore-Penrose inverse (pseudo-inverse), the divergence is

$$\operatorname{div} Y(v_i) = \sum_{j,\{v_i,v_j\}\in E} \omega_{ij} Y_{ij},$$

and

$$\Delta_0 = \begin{cases} \sum_j \omega_{ij}, \text{if } i = j \\ -\omega_{ij}, \text{if } v_i v_j \in E \\ 0, \text{otherwise.} \end{cases}$$

5.6 Practical Example

Suppose there are five students, A, B, C, D, and E, we want to rank their homework and know which is the best. Each time teacher compares two homework, and gives his preference by scores:

- A is better than B by 2 points,

- A is better than E by 8 points,

- B is better than C by 6 points,

- B is better than D by 1 point,

- C is better than A by 2 points, and

- D is better than E by 6 points.

Then we can have the pairwise comparison graph.

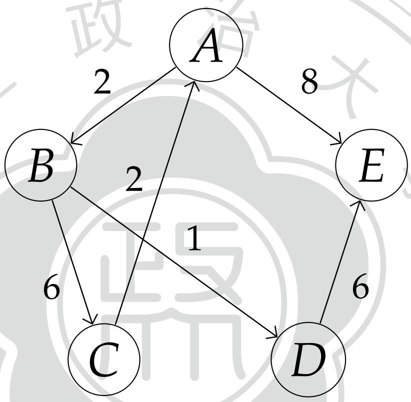

Figure 5.4: Pairwise comparison result.

We can also use the skew-symmetric matrix to represent the relations.

$$Y = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & -2 & 2 & 0 & -8 \\ 2 & 0 & -6 & -1 & 0 \\ -2 & 6 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -6 \\ 8 & 0 & 0 & 6 & 0 \end{pmatrix} \end{array}.$$

The weight matrix is aaand the

$$\operatorname{div} Y = \begin{pmatrix} -8 \\ -5 \\ 4 \\ -5 \\ 14 \end{pmatrix}.$$

We can get the matrix

$$\Delta_0 = \begin{pmatrix} 3 & -1 & -1 & 0 & -1 \\ -1 & 3 & -1 & -1 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & -1 & 0 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{pmatrix}.$$

Then

$$\Delta_0^{\dagger} = \begin{pmatrix} 0.500 & 0.318 & 0.409 & 0.045 & -0.227 \\ 0.500 & 0.864 & 0.682 & 0.409 & -0.045 \\ 0.500 & 0.591 & 1.045 & 0.227 & -0.136 \\ 0.500 & 0.682 & 0.591 & 0.955 & 0.227 \\ 0.500 & 0.500 & 0.500 & 0.500 & 0.500 \end{pmatrix}.$$

We can get the score function

$$s = -\Delta_0^{\dagger} \operatorname{div} = \begin{pmatrix} 7.36 \\ 8.27 \\ 5.81 \\ 6.64 \\ 0 \end{pmatrix}$$

And we can get the consistent skew-symmetric matrix

$$X = \begin{pmatrix} 0 & 8.27-7.36 & 5.81-7.36 & 6.64-7.36 & 0-7.36 \\ 7.36-8.27 & 0 & 5.81-8.27 & 6.64-8.27 & 0-8.27 \\ 7.36-5.81 & 8.27-5.81 & 0 & 6.64-5.81 & 0-5.81 \\ 7.36-6.64 & 8.27-6.64 & 5.81-6.64 & 0 & 0-6.64 \\ 7.36-0 & 8.27-0 & 5.81-0 & 6.64-0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0.91 & -1.55 & -0.72 & -7.36 \\ -0.91 & 0 & -2.46 & -1.63 & -8.27 \\ 1.55 & 2.46 & 0 & 0.83 & -5.81 \\ 0.72 & 1.63 & -0.83 & 0 & -6.64 \\ 7.36 & 8.27 & 5.81 & 6.64 & 0 \end{pmatrix}.$$

Chapter 6

Conclusion

In this thesis, we primarily show the elementary linear algebra that is used to survey HodgeRank and deduce combinatorial Hodge theorem. In the original paper, Jiang, Lim, Yao and Ye, proposed that HodgeRank is a suitable method for ranking datasets which are incomplete or imbalanced. In other words, most values are missing or some candidates are popular that they received a great quantity of scores. These difficulties usually happened in the modern ranking problems. We use HodgeRank to deal with these troubles and each time we only compare two candidate. Then we can get a global ranking so that we can give every candidate a score.

Pairwise comparison is practical in today's ranking problems. For example, the teacher uses MOOCs for the course this semester, and students will hand in their homework on the internet. Teacher use Peer assessment that students will give scores to a few classmates' homework after the homework deadline. Online courses is popular and have more people to sign up now. It is not possible for a teacher to score all homework. In the end of the course, teacher can use HodgeRank and obtain the ranking of all students' performance.

There are a lot of applications of HodgeRank still can try : combine HodgeRank with other method to ranking or think about how to rank when adding a new candidate or there is a candidate have no scores.

Also in this thesis, we mention that there is some little problem when we define the inner product. So in the following research, we can try to give more constraints to find the inner product space. To let the method be more perfect.

Bibliography

[1] Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346, 1950.

[2] Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika*, 39:324–345, 1952.

[3] Corinna Cortes, Mehryar Mohri, and Ashish Rastogi. Magnitude-preserving ranking algorithms. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pages 169–176, New York, NY, USA, 2007. ACM.

[4] H. A. David. *The method of paired comparisons*, volume 41 of *Griffin's Statistical Monographs & Courses*. Charles Griffin & Co., Ltd., London; The Clarendon Press, Oxford University Press, New York, second edition, 1988.

[5] Dorit S. Hochbaum and Asaf Levin. Methodologies and algorithms for group-rankings decision. *Manage. Sci.*, 52(9):1394–1408, September 2006.

[6] Xiaoye Jiang, Lek-Heng Lim, Yuan Yao, and Yinyu Ye. Statistical ranking and combinatorial Hodge theory. *Math. Program.*, 127(1, Ser. B):203–244, 2011.

[7] M. G. Kendall and B. Babington Smith. On the method of paired comparisons. *Biometrika*, 31:324–345, 1940.

[8] George Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information, 1956. One of the 100 most influential papers in cognitive science: http://cogsci.umn.edu/millennium/final.html.

[9] Frederick Mosteller. Remarks on the method of paired comparisons: I. the least squares solution assuming equal standard deviations and equal correlations. *Psychometrika*, 16(1): 3–9, 1951.

[10] Frederick Mosteller. Remarks on the method of paired comparisons: Ii. the effect of an aberrant standard deviation when equal standard deviations and equal correlations are assumed. *Psychometrika*, 16(2):203–206, 1951.

[11] Frederick Mosteller. Remarks on the method of paired comparisons: Iii. a test of significance for paired comparisons when equal standard deviations and equal correlations are assumed. *Psychometrika*, 16(2):207–218, 1951.

[12] Gottfried E. Noether. Remarks about a paired comparison model. *Psychometrika*, 25:357–367, 1960.

[13] Donald G. Saari and Vincent R. Merlin. A geometric examination of Kemeny's rule. *Soc. Choice Welf.*, 17(3):403–438, 2000.

[14] Thomas L. Saaty. A scaling method for priorities in hierarchical structures. *J. Mathematical Psychology*, 15(3):234–281, 1977.

[15] Claire Voisin. *Hodge theory and complex algebraic geometry. I*, volume 76 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, english edition, 2007. Translated from the French by Leila Schneps.

[16] Claire Voisin. *Hodge theory and complex algebraic geometry. II*, volume 77 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, english edition, 2007. Translated from the French by Leila Schneps.

[17] Douglas B. West. *Introduction to graph theory*. Prentice Hall, Inc., Upper Saddle River, NJ, 1996.